

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 31, 2005, 18:02:30 ; Search time 148.167 Seconds
(without alignments)
54.366 Million cell updates/sec

Title: US-10-083-768-12
Perfect score: 85
Sequence: 1 CADGPTLEWISFC 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_spot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 53 | 62.4 | 297 | 2 Q7UQE4 | Q7UQE4 rhodopirell |
| 2 | 50.5 | 59.4 | 387 | 2 Q98A97 | Q98A97 rhizobium l |
| 3 | 50.5 | 59.4 | 389 | 2 Q8KJF9 | Q8KJF9 rhizobium l |
| 4 | 48 | 56.5 | 1123 | 2 Q7QC63 | Q7QC63 anopheles g |
| 5 | 47.5 | 55.9 | 283 | 2 Q82CW2 | Q82CW2 streptomyce |
| 6 | 46 | 54.1 | 319 | 2 Q8RKM5 | Q8RKM5 streptomyce |
| 7 | 46 | 54.1 | 347 | 2 Q7PPU6 | Q7PPU6 anopheles g |
| 8 | 45 | 52.9 | 108 | 2 Q7RUAS | Q7RUAS neurospora |
| 9 | 45 | 52.9 | 173 | 2 Q8C4M6 | Q8C4M6 mus musculu |
| 10 | 45 | 52.9 | 209 | 2 Q6N1X5 | Q6N1X5 rhodopseudo |
| 11 | 45 | 52.9 | 209 | 2 CAE29718 | CAE29718 rhodopneu |
| 12 | 45 | 52.9 | 309 | 2 Q8XZM5 | Q8XZM5 raistonia s |
| 13 | 45 | 52.9 | 443 | 2 Q6P858 | Q6P858 phaeosphaer |
| 14 | 44 | 51.8 | 173 | 2 Q6QHD2 | Q6QHD2 gallid herp |
| 15 | 44 | 51.8 | 173 | 2 AAS48543 | AAS48543 gallid he |
| 16 | 44 | 51.8 | 178 | 2 Q6PL14 | Q6PL14 gallid herp |
| 17 | 44 | 51.8 | 178 | 2 AAT09767 | AAT09767 gallid he |
| 18 | 44 | 51.8 | 209 | 2 Q9L0S9 | Q9L0S9 streptomyce |
| 19 | 44 | 51.8 | 292 | 2 Q67642 | Q67642 gallid herp |
| 20 | 44 | 51.8 | 298 | 2 Q86653 | Q86653 gallid herp |
| 21 | 44 | 51.8 | 974 | 1 PHS2_SOLTU | PHS2_SOLTU |
| 22 | 44 | 51.8 | 997 | 2 Q6BI26 | Q6BI26 debaryomyce |
| 23 | 44 | 51.8 | 1008 | 2 Q8AY37 | Q8AY37 fundulus he |
| 24 | 44 | 51.8 | 1011 | 2 Q6VYM7 | Q6VYM7 oncorhynch |
| 25 | 44 | 51.8 | 1011 | 2 AAQ82787 | AAQ82787 |
| 26 | 44 | 51.8 | 1022 | 1 A7IA_TORCA | A7IA_TORCA |
| 27 | 44 | 51.8 | 1023 | 1 A1A1_HUMAN | A1A1_HUMAN |
| 28 | 44 | 51.8 | 1025 | 2 Q7ZYK8 | Q7ZYK8 xenopus lae |
| 29 | 44 | 51.8 | 1028 | 1 A1A4_RAT | A1A4_RAT |
| 30 | 43.5 | 51.2 | 405 | 2 Q9KIE9 | Q9KIE9 streptomyce |
| 31 | 43.5 | 51.2 | 934 | 2 Q9NEX6 | Q9NEX6 caenorhabdi |

32 43 50.6 127 2 Q9N0Z5
33 43 50.6 171 2 Q8HYW6
34 43 50.6 176 2 Q866A9
35 43 50.6 245 2 Q9M060
36 43 50.6 407 2 Q6NMU4
37 43 50.6 407 2 Q9VK55
38 43 50.6 407 2 AAS15697
39 43 50.6 469 2 Q37839
40 43 50.6 490 2 Q04270
41 43 50.6 509 2 Q8BIG9
42 43 50.6 960 2 Q80U28
43 43 50.6 962 2 Q9IYI9
44 43 50.6 1000 2 Q724I9
45 43 50.6 1009 2 Q98SL3

Q9n0z5 oryctolagus
Q8hyw6 bos taurus
Q866a9 equus cabal
Q9m060 arabidopsis
Q6nmu4 drosophila
Q9vk55 drosophila
Aas15697 drosophil
Q37839 bacterioph
Q04270 chlamydomon
Q8big9 mus musculu
Q80u28 mus musculu
Q9iyi9 mus musculu
Q724i9 homo sapien
Q98sl3 electrophor

ALIGNMENTS

RESULT 1
Q7UQE4
ID Q7UQE4 PRELIMINARY; PRT; 297 AA.
AC Q7UQE4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RB6375;
OS Rhodopirellula ballica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
RL EMBL; BX294144; CAD74759.1;
DR InterPro; IPR00194; ATPase_a/bcentre.
DR InterPro; IPR003169; GYF.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50823; GYF; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31805 MW; 475P670F02C78E9B CRC64;

Query Match 62.4%; Score 53; DB 2; Length 297;
Best Local Similarity 72.7%; Pred. No. 1.1;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ADGPTLEWIS 12
| | | | | : : : : :
Db 175 ADGPTMKQWIS 185

RESULT 2
Q98A97
ID Q98A97 PRELIMINARY; PRT; 387 AA.
AC Q98A97;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mlr6096 protein.
GN OrderedLocusNames=mlr6096;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=WAFF303099;
RA MEDLINE=21082936; PubMed=11214974;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti (supplement).";
RL DNA Res. 7:381-406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WAFF303099;
RA MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003008; BAB5240.1; -.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR003673; CAIB_BAIF.
DR Pfam: PF02515; CoA_transf_3; 1.
KW Complete proteome.
SQ SEQUENCE 387 AA; 42226 MW; 64643BEC8F25518 CRC64;

Query Match 59.4%; Score 50.5; DB 2; Length 387;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Qy 1 CADGPTL-----REWISFC 14
|||||:
Db 237 CADGKEVIFSQNDREWVNF 257
|||||:

RESULT 3
ID Q8KJF9 PRELIMINARY; PRT; 389 AA.
AC Q8KJF9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PUTATIVE RACEMASE/DEHYDRATASE PROTEIN.
GN Name=mel181;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R7A;
RA MEDLINE=21999272; PubMed=12003951; Cruickshank R.W., Gouy J.,
RA Sullivan J.T., Trzebiatowski J.R., Fleetwood D.J., McCallum N.G., Rossbach U.,
RA Brown S.D., Elliot R.M., Fleetwood D.J., Webb R.J., Ronson C.W.;
RA Stuart G.S., Weaver J.E., Webb R.J., de Bruijn F.J., Ronson C.W.;
RT "Comparative sequence analysis of the symbiosis island of
RT Mesorhizobium loti strain R7A.";
RL J. Bacteriol. 184:3086-3095(2002).
DR EMBL: AL672113; CAD31586.1; -.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR003673; CAIB_BAIF.
DR Pfam: PF02515; CoA_transf_3; 1.
SQ SEQUENCE 389 AA; 42703 MW; 6678D2C96A7E5204 CRC64;

Query Match 59.4%; Score 50.5; DB 2; Length 389;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Qy 1 CADGPTL-----REWISFC 14
|||||:

```

```

Db 243 CADGKEVIFSQNDREWVNF 263
|||||:
|||||:

RESULT 4
ID Q7QC63 PRELIMINARY; PRT; 1123 AA.
AC Q7QC63;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP1221.
GN Name=agGS3078; ORFNames=ENSGG00000018866;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -|- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB0100859; EAA08177.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008998; F:homocysteine S-methyltransferase activity; IEA.
DR GO: GO:0004672; F:protein kinase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR003726; S_methyl_trans.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF02574; S-methyl_trans; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.
SQ SEQUENCE 1123 AA; 120006 MW; D3CC001D8D4882AF CRC64;

Query Match 56.5%; Score 48; DB 2; Length 1123;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ADGPTLREWISF 13
|||||:
Db 969 ADHPTVRFWISF 980
|||||:

RESULT 5
ID Q82CW2 PRELIMINARY; PRT; 283 AA.
AC Q82CW2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative IclR-family transcriptional regulator.
GN OrderedLocusNames=SAV5226;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RA MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RL microorganism Streptomyces avermectilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AF005042; BAC72938.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR005471; HTH_IcLR.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF01614; ICLR; 1.
KW Complete proteome.
SQ SEQUENCE 283 AA; 30503 MW; F63B1705578EE67 CRC64;

Query Match 55.9%; Score 47.5; DB 2; Length 283;
Best Local Similarity 50.0%; Pred. No. 9.4;
Matches 8; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 CADGPT---LREWISF 13
||| : ||| : |||
152 CAEGPTTAVHEWVDF 167

DB 152 CAEGPTTAVHEWVDF 167

RESULT 6
Q9RKM5 PRELIMINARY; PRT; 319 AA.
ID Q9RKM5
AC Q9RKM5
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative MerR family transcriptional regulator.
GN Ordered locus names=SC04102; ORFNames=SCD17.06c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.P., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
RL EMBL; AL939118; CAB56383.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR InterPro; IPR009061; Putativ_DNA_bnd.
DR Pfam; PF00376; MerR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MERR; 1.
DR PROSITE; PS0937; HTH_MERR; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 319 AA; 34841 MW; 1F51905A8BA5365E CRC64;

Query Match 54.1%; Score 46; DB 2; Length 319;
Best Local Similarity 70.0%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 DGPTLREWIS 12

```

```

DB 258 DGPELREWLA 267
||| ||||| :
||| ||||| :

RESULT 7
Q7PPP6 PRELIMINARY; PRT; 347 AA.
ID Q7PPP6
AC Q7PPP6
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000020769 (fragment).
GN Name=ENSANGG0000018280;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008944; EAA10075.2; -.
DR GO; GO:0008898; F:homocysteine S-methyltransferase activity; IEA.
DR InterPro; IPR003726; S_methyl_trans.
DR Pfam; PF02574; S-methyl_trans; 1.
FT NON TER 1
SQ SEQUENCE 347 AA; 38585 MW; 66FF58A100CDA4F CRC64;

Query Match 54.1%; Score 46; DB 2; Length 347;
Best Local Similarity 61.5%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 CADGPTLREWISF 13
||| : ||| : |||
201 CDEYTVRFWISF 213

DB 201 CDEYTVRFWISF 213

RESULT 8
Q7RUA5 PRELIMINARY; PRT; 108 AA.
ID Q7RUA5
AC Q7RUA5
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein B24B19.30.
GN Name=NCU03933.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Koche G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysellis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen M., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nuebaum C., Birren B.;
RT "The genome sequence of the filamentous fungus Neurospora crassa.";
RL Nature 0.0-0(2003).

```

CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AABX01000719; EAA28336.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 108 AA; 11994 MW; 093DC0D9617A252E CRC64;

Query Match 52.9%; Score 45; DB 2; Length 108;
 Best Local Similarity 50.0%; Pred. No. 9.1;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CADGPTLREWISFC 14
 Db 70 COCQPILRNLSWC 83

RESULT 9

Q8C4M6 ID Q8C4M6 PRELIMINARY; PRT; 173 AA.

AC Q8C4M6
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
 DE library, clone: C130070D15 product: unclassifiable, full insert
 DE sequence.
 GN Name: C130070B15Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1] TaxID=10090;

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 [2]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN PANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 [3]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
 RA The PANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 [4]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 [5]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 [6]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK081706; BAC38302.1; -.
 DR MGD; MGI:2444974; C130070B15Rik.

SQ SEQUENCE 173 AA; 19340 MW; 6227DD6725E52FCD CRC64;

Query Match 52.9%; Score 45; DB 2; Length 173;
 Best Local Similarity 63.6%; Pred. No. 15;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 GPTLREWISFC 14

Db 75 GVTMREWASWC 85

RESULT 10

Q6NX5 ID Q6NX5 PRELIMINARY; PRT; 209 AA.

AC Q6NX5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=RP44277;
 OS Rhodopseudomonas palustris.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Rhodopseudomonas.
 OX NCBI_TaxID=1076;
 [1]

SEQUENCE FROM N.A.

RC STRAIN=CGA009 / ATCC BAA-98;
 RX PubMed=14704707; DOI=10.1038/nbt923;
 RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
 RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
 RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
 RA Harrison F.H., Gibson J., Harwood C.S.;
 RT "Complete genome sequence of the metabolically versatile
 photosynthetic bacterium Rhodopseudomonas palustris."
 RL Nat. Biotechnol. 22:55-61(2004).
 DR EMBL; BX572606; CAE29718.1; -.
 DR InterPro; IPR008938; ARM.

Complete proteome; Hypothetical protein.

SQ SEQUENCE 209 AA; 23238 MW; 6FE082A84BD040EE CRC64;

Query Match 52.9%; Score 45; DB 2; Length 209;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

Qy 1 CADG-----PTLREWIS 12

Db 98 CADTGVEALPTIREWLS 115

RESULT 11

CAE29718 PRELIMINARY; PRT; 209 AA.

ID CAE29718
 AC CAE29718;

```

DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN RPA4277.
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RA "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572606; CAE29718.1; -.
KW Hypothetical protein.
SQ SEQUENCE 209 AA; 23238 MW; 6FE082A84DB040EE CRC64;

Query Match 52.9%; Score 45; DB 2; Length 209;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 1; Indels 6; Gaps 1;

QY 1 CADG-----PTLREWIS 12
   ||| ||| ||| |||
DB 98 CADTGYEAALPTIREWLS 115

RESULT 12
Q8XZN5
ID Q8XZN5 PRELIMINARY; PRT; 309 AA.
AC Q8XZN5;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE PROBABLE TRANSCRIPTION REGULATOR PROTEIN.
GN Name=RS04642; OrderedLocusNames=RS1360;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646064; CAD15062.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PROSITE; PS00931; HTH_LysR; 1.
KW Complete proteome.
SQ SEQUENCE 309 AA; 33774 MW; 733551741CE83182 CRC64;

Query Match 52.9%; Score 45; DB 2; Length 309;
Best Local Similarity 70.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 CADGPTLREW 10
   ||| ||| |||
DB 221 CTDCGAVLREW 230

RESULT 13
Q9P858
ID Q9P858 PRELIMINARY; PRT; 443 AA.
AC Q9P858;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Phaeosphaeria nodorum (Septoria nodorum).
OG Plasmid plsal.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Phaeosphaeriaceae; Phaeosphaeria.
OX NCBI_TaxID=13684;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BS444;
RA Rawson J.M.;
RT "Transposable elements in the phytopathogenic fungus Stagonospora
RL Thesis (2000), PhD thesis, University of Birmingham, UK.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BS444;
RA Rawson J.M., Cutler S.B., Caten C.E.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277966; CAB91876.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 443 AA; 49466 MW; 367E0762EE839E68 CRC64;

Query Match 52.9%; Score 45; DB 2; Length 443;
Best Local Similarity 58.3%; Pred. No. 42;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 CADGPTLREWIS 12
   ||| ||| ||| |||
DB 170 CSENGTLREWIT 181

RESULT 14
Q6QHD2
ID Q6QHD2 PRELIMINARY; PRT; 173 AA.
AC Q6QHD2;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE P32 (Fragment).
OS Gallid herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
OX NCBI_TaxID=10386;
RN [1]
RP SEQUENCE FROM N.A.
RA Villarreal L.Y., Brandao P.E., Ferreira A.P., Dorretto L.J.,
RA D'elboux A.N.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY541676; AAS48543.1; -.
DR InterPro; IPR007110; IG-like.
FT NON_TER 1
FT NON_TER 173
SQ SEQUENCE 173 AA; 19130 MW; 5AE4A1956CEB9B13 CRC64;

Query Match 51.8%; Score 44; DB 2; Length 173;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CADGPTLREWISFC 14
   ||| ||| ||| |||

```

```
DB      120 CLDMPPLRPWTVC 133

RESULT 15
AAS48543
ID AAS48543 PRELIMINARY; PRT; 173 AA.
AC AAS48543;
DT 10-MAR-2004 (TREMELrel. 27, Created)
DT 10-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 10-MAR-2004 (TREMELrel. 27, Last annotation update)
DE F32 (Fragment).
OS Gallid herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
OX NCBI_TaxID=10386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP01;
RA Villarreal L.Y., Brandao P.E., Ferreira A.P., Doretto L.J.,
RA D'elboux A.N.;
RT "Molecular epidemiology of an outbreak of infectious laryngotracheitis
RT in Brazilian poultry.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY541676; AAS48543.1; -.
FT NON_TER 1
FT NON_TER 173
SQ SEQUENCE 173 AA; 19130 MW; 5AE4A1956CEB9B13 CRC64;

Query Match 51.8%; Score 44; DB 2; Length 173;
Best Local Similarity 50.0%; Pred. NO. 23;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1 CADGPTLRWISFC 14
DB      120 CLDMPPLRPWTVC 133

Search completed: January 31, 2005, 18:21:43
Job time : 151.167 secs
```